

Taller 1

Nicolás Chávez

9 de septiembre 2020

1 Presentación

Mi nombre es Nicolás Chávez, tengo 21 años y vivo en Bogotá. Estudio Economía en la Universidad Nacional de Colombia en décimo semestre. Siempre me he interesado mucho por los temas sociales y por intentar lograr una sociedad más justa y mejor para todas las personas. Esta fue una de mis motivaciones para estudiar economía, ya que en general lo que percibía era que aquellos que formulan y ejecutan política pública son economistas y yo quería poder entender y llegar a generar políticas que mejoraran la sociedad. Fue así que decidí estudiar economía, ya que tenía la convicción de que esta carrera podría ayudarme a cumplir esos objetivos de transformación. Durante la carrera también me di cuenta que la economía es fundamental para entender como funciona el mundo, las relaciones de poder, el mercado y en general como funcionan muchos aspectos de la sociedad.

Debido a mi interés de hacer política pública y poder hacer cambios, es importante aprender a diseñar estas políticas y tener instrumentos y herramientas para poder evaluar el impacto que podrían tener estas políticas. Además también es importante este conocimiento para evaluar políticas anteriores y poder saber cuales funcionan y cuales no. Es por esto que me interesó tomar el curso ya que espero aprender a evaluar políticas públicas y así poder aumentar mis capacidades como profesional para aplicar todos estos conocimientos. También es importante para repasar y fortalecer mis conocimientos en econometría y poder ver una aplicación practica de esta rama de la economía tan útil para la investigación académica y para la vida profesional.

2 Gentzkow and Shapiro

El paper de Gentzkow y Shapiro da una serie de recomendaciones a la hora de hacer trabajos académicos de carácter empírico en las ciencias sociales (ciencia política, economía, psicología, entre otras), los cuales permiten tener mayor orden y claridad, además de facilitar el manejo de datos y de código a la hora de necesitarlos una vez más para replicarlos o modificarlos.

En el primer capítulo, se preguntan sobre ¿qué es hacer ciencias sociales empíricas?, a lo que responde tres conceptos cruciales: primero es hacer bue-

nas preguntas, segundo buscar los datos relevantes, tercero, diseñar un análisis estadístico, y por último, escribir los resultados. Por esto, va a presentar una serie de elementos claves para optimizar la realización de estos cuatro pasos y sobre todo como poder trabajar en equipos de investigación de mejor forma.

En el segundo capítulo trata sobre la automatización, en la cual presentan dos reglas principales. Primer automatizar todo lo que se pueda automatizar, segundo, escribir un solo script que ejecute todo el proceso de principio a fin. Esto para tener una mayor facilidad a la hora de replicar y una mayor eficiencia a la hora de hacer otras regresiones con otros datos para el mismo trabajo.

El tercer capítulo habla del control de versiones, en donde expone el problema de tener múltiples versiones con muchas modificaciones de diversos coautores de un mismo documento o código. Para esto recomiendan tener los archivos y datos en versiones de control en donde se puedan ver los cambios que se han hecho y quienes los han hecho, para así poder deshacerlos si es necesario y llevar un control efectivo de quien realiza los cambios (herramientas como Google Docs son útiles para esto).

El cuarto capítulo trata sobre los directorios. Para organizar los directorios es importante separa los directorios según su función, separar los archivos de acuerdo a si son de entrada (como bases de datos) o de salida como pueden ser los resultados y por último hacerlo portable, para que se pueda manejar en diferentes equipos o en la nube.

En el quinto capítulo habla sobre las claves. Una clave es una variable o grupo de variables que únicamente identifica los elementos de una tabla. Para el manejo de las claves recomiendan almacenar los datos "limpios" en tablas únicas con claves que no se confundan y mantener los datos normalizados tanto como la línea de código lo permita.

El quinto capítulo trata sobre la abstracción, que se refiere poder pasar de lo específico a lo general, pero esta vez refiriéndose al código. Lo que trata de decir es poder abstraer el código de las especificidades de los datos a una forma más general para poder utilizarlo con datos diferentes sin necesidad de repetirlo o que se generen errores por copiar y pegar, y no modificar partes del código porque tenían los datos anteriores. Esto es recomendable siempre que aumente la claridad y elimine redundancia; pero si el código no se va a utilizar más o es muy específico, esto puede generar trabajo extra innecesario.

En el capítulo siete trata sobre la documentación, sobre todo a la hora de comentar el código para dar claridad y poner referencias sobre la razón de los datos y los procedimientos estadísticos. Aquí es importante tener en cuenta que no hay que escribir documentación que no se vaya a mantener en el script final, y que hay que procurar que el código sea auto documentado, es decir que mediante el código demos a entender lo que hay que hacer o lo que no se puede hacer para evitar un exceso de comentarios y dar más claridad.

Por último, da unas recomendaciones sobre el manejo de tareas cuando se trabaja en grupos y del estilo del código. Es recomendable trabajar con un sistema de manejo de tareas que optimice la comunicación y pueda llevar un control adecuado sobre los avances y las tareas. Por otro lado aclaran que el correo electrónico no es un sistema de manejo de tareas.